REMARKS/ARGUMENTS

In the specification, pages 1 to 9 have been amended by the insertion of Section Headings. Consequential changes have occurred to the pagination of pages 1 to 21.

Claim 27 has been amended to claim a computer readable medium.

Claim 1 has been amended to distinguish more clearly between interactive data to be associated with an object and data representative of the object which are combined together in a data sequence, and corresponding amendments have been made to claim 14.

Claims 1-5, 11-18, 21 and 24-28 are rejected under 35 U.S.C. 103(a) as being unpatentable over US 5,590,262 (Isadore-Barreca) in view of US 5,708,845 (Wistendahl et al).

However, we submit that Isadore-Barreca teaches the use of a video as a computer interface (e.g. col. 3 lines 64-65, vol. 4 line 25). An edit list, indicating when an object enters or leaves a scene, is used to define a sequence (col. 6 line 33-61, col. 7 lines 6-10 and 12-13). A key-frame is defined at least at the beginning of the sequence, and, dependent on the movement of the camera and objects within the sequence, and the length of the sequence, additional key-frames may be defined within and at the end of the sequence (col. 7 line 10 and col. 7 line 31 – col. 8 line 55). Hotspots, corresponding to the position of objects, are manually allocated in the key-frames (col. 4 line 39, col. 7 line 11 to col. 8 line 55, col. 8 lines 61-62, col. 9 lines 51-67). In use, a user stops the video at any frame (col. 4 lines 41-42, col. 8 line 56). The nearest previous key-frame is displayed (col. 8 line 61 – col. 9 line 8, col. 10 lines 40-41, col. 12 lines 17-20, 54-56). The user selects a hotspot in the key-frame to access associated information etc. (col. 10 lines 38-55).

There is no teaching or suggestion in Isadore-Barreca of "parsing [a] video program by identifying separate shots in the video program to produce an edit list", or of "identifying shots containing related content to form a sequence of shots containing related content" as claimed in claim 1, but only of using a pre-existing edit list (col. 7 lines 12-16) or of using a "rough equivalent" of an edit list by [manually] scanning the video and noting entrance and exit points of objects (col. 7 lines 16-24). That is, parsing as disclosed in the present invention implies some form of analysis of the input, i.e. the video program, in order to determine its structure. However, there is no teaching in Isadore-Barreca of automatic analysis of the input. The shots are identified only by importing an EDL file or roughly assigning them by operator intervention. Thus, the teaching of the present invention is that shots containing related content are identified to form a sequence of shots, as claimed in claim 1, whereas in contrast in Isadore-Barreca the "sequence" consists only of contiguous frames containing a same object, that is, a sequence of frames between the entrance and exit of the object. Thus, in the present invention, "the sequences consists of a series of semantically related shots and, for example, one sequence may contain all the shots that feature [a] lead singer in a pop group" (paragraph bridging pages 11 and 12). There is no teaching in Isadore-Barreca of forming a sequence of non-consecutive shots containing a same object as in the present invention. Thus there is no teaching in Isadore-Barreca of a sequence of shots containing related content. A "sequence" in Isadore-Barreca is merely a series of sequential frames separated from a next "sequence" by transitions. Thus a "sequence" in Isadore-Barreca is a shot; the in and out points define shot boundaries. There is no teaching of a sequence, as disclosed in the present invention, corresponding to a scene or series of shots with related content.

34

There is no suggestion in Isadore-Barreca of selecting attributes of objects, such as shape, size, position, color, texture and intensity gradient (page 14 last full paragraph) "for tracking the object through the sequence of shots" as claimed in claim 1. The mere selection of coordinates of a hot spot containing an object taught by Isadore-Barreca (col. 10 lines 35-38) is not comparable to selecting attributes with which an object can be tracked through a sequence of different shots in which the position of the object is likely to be constantly changing in a disconnected manner between shots, so that knowledge of location alone of an object in one shot is of little or no assistance in identifying and tracking the object in a subsequent shot, as is done in the present invention using in addition color, shape, texture etc.

Moreover, there is no suggestion in Isadore-Barreca of embedding (multiplexing) (paragraph bridging pages 10 and 11) interactive data with data representative of an object in a single data sequence as described in the present invention and claimed in claim 1. Instead, in Isadore-Barreca a pointer is used from the key-frame to data stored in a database, since in Isadore-Barreca the video is used merely as a computer interface (col. 9 lines 16-50).

As conceded in the Office Action, there is no disclosure in Isadore-Barreca of tracking an object throughout a video as in the present invention, because there is no incentive in Isadore-Barreca to identify the same object in different shots. The only requirement in Isadore-Barreca is to nominate at least one key-frame in each shot and nominate at least one hotspot in the at least one key-frame. Contrary to the contention of the Office Action, since hotspots are identified only in key-frames in Isadore-Barreca, there is no requirement to draw an outline "around the object in each frame of the sequence" in Isadore-Barreca, and therefore no requirement to track objects through a shot or between shots.

We therefore submit that Isadore-Barreca teaches a method of creating interactive multimedia works (col. 1 line 10) by converting a conventional audio/video work into a practical computer user interface (col. 3 lines 51-65). An edit decision list is created during a final creation of the video work including frame identifier codes (col. 6 lines 35-50; col. 7 lines 8-10). Where no edit decision list exists a rough equivalent can be made by scanning the video and notating the in points and out points of the scenes (col. 7 lines 16-28). The edit decision list is used to define a sequence and a key-frame is nominated in each sequence (col. 7 lines 33-35). The key-frames are defined in the conventional video work (col. 6 lines 14-15) and a record made of the SMPTE or other identifying code of the frame (col. 6 lines -16-24). A key-frame database is created (col. 6 lines 25-32; col. 7 lines 31-35). Items within the key-frames about which it is desired to provide information are defined as objects (col. 9 lines 8-34; col. 9 line 51 – col. 10 line 55) and marked with an identification overlay. A user may stop the video at any frame and select an object (col. 4 lines 25-42) from an associated key-frame (col. 8 line 56 – col. 9 line 8).

We therefore submit that there is no disclosure in Isadore-Barreca of at least the following elements claimed in claim 1:

a)   means for parsing the video program by identifying separate shots in the video program to produce an edit list (in Isadore-Barreca either a pre-existing edit list is used or a rough equivalent is produced manually)

b)   means for identifying shots containing related content to form a sequence of shots;

c)   means for extracting attributes (color, shape, texture etc.) of objects; and

36

d)   tracking means for using the attributes of an object for tracking the object through the sequence of shots.

Wistendahl et al discloses an interactive digital media program for performing a function when a user selects an object on a video display without embedding any data in the video stream (col. 2 lines 35-58). Thus location coordinates and frame addresses of mapped objects are maintained separately from the media content (col. 2 lines 59-62, col. 4 line 65 – col. 5 line 1). In order to index objects in a video the objects are encircled and their location and frame number stored (col. 10 lines 1-10). The procedure has to be repeated for all objects to be indexed in the frame and for all frames of the video (col. 10 lines 10-12), unless an object is stationary with respect to the camera, in which case the same data is saved for all frames or for the first and last frame in which the object is unchanged (col. 10 lines 15-26). A motion estimation tracking tool may be used to track a moving object through sequential, successive frames from a key-frame to a last frame in a shot in which the object is detected (col. 11 lines 3-44).

Wistendahl et al therefore does not teach "means for parsing the video program by identifying separate shots in the video program to produce an edit list" or "means for identifying shots containing related content to form a sequence of shots containing related content" as claimed in claim 1. Wistenthal et al does not suggest or hint at tracking an object through a sequence of different shots in which the position of the object is likely to be constantly changing by recognizing an object in different shots from stored attributes of the object. Clearly, the location of the object, which is all that is stored in the cited passage (col. 11 lines 41-44) is not sufficient as an attribute for this task as described and claimed in the present invention. Moreover, Wistendahl et al teaches away from "embedding interactive content data with data representative of said object in a data sequence" as claimed in claim 1.

37

Therefore, since parsing is not disclosed in Isadore-Barreca or Wistendahl et al no combination of Isadore-Barreca and Wistendahl et al would result in parsing for creating an edit list as disclosed in, for example the paragraph bridging pages 11 and 12 and claimed in claim 1.

As indicated above, there is no incentive in Isadore-Barreca for tracking an object through successive frames, since only objects in key-frames are indexed. In use, when the user attempts to stop the video on any frame the most recent key-frame is displayed instead. This differs fundamentally from the present invention in which an object is tagged in substantially all frames in which the object appears. Thus, while Isadore-Barreca associates added data with selected key-frames, the present invention associates added data with video objects wherever they appear. Moreover, even if the motion tracking of objects moving through a shot of Wistendahl et al were applied to the disclosure of Isadore-Barreca, it would not result in the advantage of the present invention of grouping together shots throughout the video so that tracking of an object can be confined to those frames likely to contain the object (paragraph bridging pages 11 and 12).

Moreover there is no suggestion in Isadore-Barreca of embedding the interactive data in a data sequence with data representative of the object, as claimed in claim 1, and Wistendahl et al positively teaches away from any such combination of video and interactive data.

Thus, we submit that Wistendahl teaches a system for using media content in an interactive digital media program (col. 1 lines 6-8) without embedding codes in the original media content (col. 2 lines 36-37; 62-65). A frame of the video is displayed on an editing subsystem and an outline drawn around an object and the pixel elements constituting the outline and the frame reference saved (col. 9 line 65 – col. 10 line 5). A hyperlinking tool is used to

define a link between the object outlines and another function to be performed (col. 10 lines 5-9). The procedure is repeated for all objects in the frame and all frames in the video (col. 10 lines 10-12). The same object outline can be used in a succeeding frame if the object is stationary (col. 10 lines 16-18). Motion tracking and motion estimating techniques may be employed for motion tracking of an unchanging object across a sequence of frames (col. 10 lines 34-42). An object is outlined in a key-frame and the outline data, position and frame address are saved (col. 11 lines 10-13). Motion tracking is used to detect the last of sequential frames in which the object is detected and the position of the object and the last frame address saved to avoid having to draw the outline around the object in each intervening frame (col. 11 lines 15-22).

Since Wistendahl does not unambiguously disclose extracting attributes of objects to track an object through a sequence of shots, and does not suggest or hint at means for parsing the program to produce an edit list or means for identifying shots containing related content, we submit that Wistendahl does not disclose some, and arguably not any, of the features of claim 1 which are not disclosed by Isadore-Barreca.

Therefore, it is submitted that no combination of Isadore-Barreca and Wistendahl et al could result in the invention claimed in claim 1.

As to claim 2, there is no suggestion in Isadore-Barreca or Wistendahl et al of producing a hierarchy of groups of shots as claimed in claim 2, in which shots are grouped into sequences by a scene grouper which compares the key-frames from each shot with key-frames from other shots using low level features such as correlelograms, data maps and textures, so that shots having similar content are grouped together into a hierarchical structure into groups of shots having a common theme in order to create a content tree to aid in the selection of objects and

39

improve the efficiency of subsequent object tracking, so that searching for a particular object is carried out only in related shots and not through all shots of the video. (paragraph bridging pages 12 and 13).

As to claim 3, Isadore-Barreca or Wistendahl et al do not disclose any means for inputting criteria for recognizing a change of shot because the only criterion used is of an object entering or leaving a frame, whereas in the present invention a variety of criteria may be input such as means for detecting shot changes, camera angle changes, wipes, dissolves and other editing function and optical transition effects and comparing edge maps (paragraph bridging pages 11 and 12).

As to claim 4, the passage cited in the Office Action appears to refer only to rectangular coordinates of a hot spot within which an object is located. We submit there is no disclosure in Isadore-Barreca or Wistendahl et al of performing edge detection of an object within a boundary and storing a geometric model of the object of which the edges are detected, as claimed in claim 4.

As to claim 5, we submit that there is no disclosure in Isadore-Barreca or Wistendahl et al of extracting at least one of the attributes listed for utilising the attributes of the object for tracking the object through a sequence of shots, as claimed in claim 5 as dependent on claim 1, Isadore-Barreca merely recording the location of an object in a key-frame but not attempting to use that location information for tracking the object through other frames or shots. In any case, there is no suggestion in Isadore-Barreca of recording "time series statistics based on said attribute" as claimed in claim 5.

As to claim 8, the passage cited in the Office Action at Wistendahl et al col. 11 lines 3-24 teaches only motion tracking of an object in a movie or video sequence. There is no suggestion of "updating the stored attributes of the object as the attributes change from time to time" as claimed in claim 8. On the contrary, Wistendahl et al suggests that only unchanging, rotating or partially occluded objects can be tracked, even with advanced techniques (col. 10 lines 48-51).

As to claims 14, 15, 16, 17, 18 and 21, corresponding arguments apply as have been presented in respect of claims 1, 3, 2, 4, 5 and 8 respectively.

The examiner has acknowledged that claims 6, 7, 9 and 10 are directed to allowable subject matter, subject only to amendment in respect of 35 U.S.C. 112, 2nd paragraph and new claims 28-31 have been formed by combining amended claim 1 with claims 6, 7, 9 and 10 respectively.

The examiner has also acknowledged that claims 19, 20, 22 and 23 are directed to allowable subject matter, and new claims 32-35 have been formed by combining original claim 14 with claims 19, 20, 22 and 23 respectively.

Applicant respectfully requests that a timely Notice of Allowance be issued in this case.

<div align="right">

Respectfully submitted,

SEYFARTH SHAW LLP

</div>

Date: 5/23/06

By

Timothy J Keefer, Reg. No. 35,567
Attorney for Applicant

55 E. Monroe St., Suite 4200
Chicago, IL 60603-5803
Telephone: (312) 346-8000
Facsimile: (312) 269-8869
e-mail: tkeefer@seyfarth.com